Intrinsic Shape Alignment of Early versus Late Type Galaxies

Undergraduate Research Thesis

Presented in partial fulfillment of the requirements for graduation with research distinction in Astronomy and Astrophysics in the undergraduate colleges of The Ohio State University

by

Dominic Flournoy

The Ohio State University May 2019

Project Advisor: Professor Barbara Ryden, Department of Astronomy

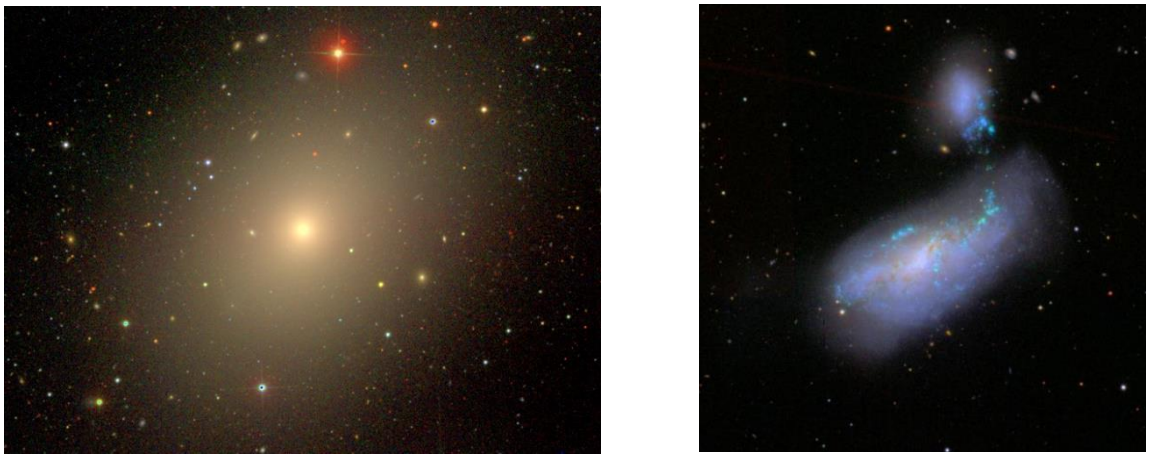
ABSTRACT

When looking at pairs of neighboring galaxies, a correlation between the alignment of the pair and their physical separation can be made. In studying this we can better understand the ways in which early and late type galaxies align themselves in the universe. Studying these pairs of closely separated galaxies allows us to make inferences on how neighboring galaxies affect each other's formation and evolution. I use 400,000 galaxies from the Sloan Digital Sky Survey (SDSS) to compare the alignment correlation function for galaxies of early and late type. Galaxies are classified as "early" if they are centrally concentrated, contain a red stellar population, and show no evidence for recent star formation. By contrast, galaxies are "late" if they are less concentrated, contain blue stars, and show spectral emission characteristic of star forming regions. The trend among galaxies is a parallel alignment until they reach a separation of 20 kiloparsecs (kpc), in which they have a perpendicular alignment after. This alignment lasts until the galaxies are separated by over 100 kpc and thereafter lose their alignment correlation. This is seen in some classifications of early type galaxies but has stronger support in all classifications of late type. These findings show there is an intrinsic alignment function for early and late type galaxies in which there is a stronger alignment for late type, meaning there is a different evolutionary formation that exists between the two types.

INTRODUCTION

In defining the difference between an early and late type galaxy, it is not enough to assume that an early galaxy is older from forming earlier in the universe and a late type is younger from forming later. The definition of early and late type galaxies is a nomenclature that stuck and since has been proven not to rely solely on age but now on other characteristics the galaxy has. This is a result of adapting a better theory of galaxy formation and learning that a late type galaxy can be older than an early type and the same vis versa. Now, early type galaxies are classified by being red in color, having a centrally concentrated light profile, and little to none star formation ongoing in the galaxy. By contrast, late type galaxies are blue in color, have an exponentially spread light profile, and show evidence of recent star formation (within about 100 million years) throughout. It is also important to note that red galaxies are not a "fire truck red" color but a lighter red hue as seen in Figure 1[1]. Because of the differences in characteristics these two types have, this has also shown to effect aspects such as evolutionary track, galaxy shape, and luminosities. Along with this, it is seen that pairs of galaxies tend to align differently with their neighbor based on whether they belong in the early or late category. The main objective in this analysis is to find the intrinsic shape alignment correlation that exists between early and late type galaxy pairs.

[1] Blanton et al. (2017).



NGC 4636: (u – r) = 2.92, z = 0.003 NGC 4490: (u – r) = 1.40, z = 0.002

Figure 1 – Examples of a red (left) and blue (right) galaxy to show the differences in color and morphology.

In conducting this analysis, there are many attributes used to begin making an alignment comparison. Using the Sloan Digital Sky Survey (SDSS), a catalogue of galaxy data is readily made available to the public and is used as the source of data throughout this paper. Using the data the SDSS provides, a complete analysis of a galaxy can be made and used in understanding the shape alignment correlation among them. The shape of the galaxy itself is found by using the axis ratio (defined as parameter q) and the position angle on the sky (parameter φ). The galaxies used in this sample need to be within a certain distance which is based off the spectroscopic redshift (a measure of the rate an object is moving away from another) the SDSS provides in which the average redshift within this sample is z ~ 0.14. The light that the SDSS telescopes capture is received through five different filters known as the *ugriz* filters. Because the r-band has the highest response out of these five filters, it is the band used when retrieving data from the SDSS. The u-band magnitude from the SDSS is only used in making the color calculations. The high response given by the r-band is useful from getting an accurate measurement of the apparent magnitude (observed brightness in the sky) to make the correct calculation of the absolute magnitude (actual brightness in the sky if a set distance away) for making comparisons in color and light profiles.

Even with the amount of wonderful work put into the data analysis for the SDSS, it can still be wrong at times. One of the biggest faults it has pertaining to this analysis is the problem of galaxy shredding. This is a result of the SDSS analysis program thinking that one galaxy is two separate galaxies. When this happens, it adds an error to the alignment correlation because a galaxy that is seemingly located within itself will have a high parallel alignment. This is mostly caused by a patch of recent star formation in the galaxy being bright enough to be read as its own separate galaxy. It can also be the result of a galaxy being seen edge-on and the central dust band splits the galactic center, making the SDSS believe this is two separate galaxies with a high parallel alignment existing between them. An example is shown in Figure 2[2] where a single galaxy is being read as two indicated by the squares representing the galactic centers. If this happens, the galaxy must be excluded from the data set of comparison as to not give one

[2] Blanton et al. (2017).

alignment more favor over the other. After these galaxies are removed from the data set, the analysis of the data set can begin without worrying about uncorrected influences in the alignment correlation.

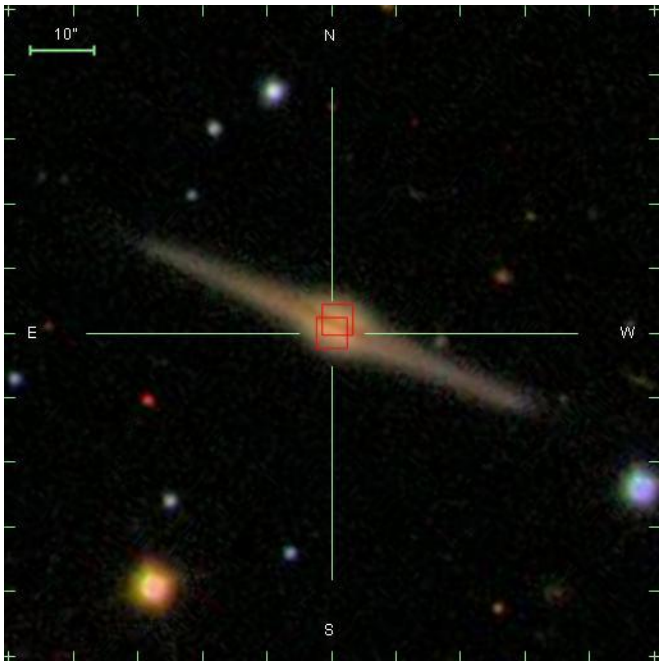


Figure 2 – An example of an edge-on spiral galaxy that has been shredded showing the two galactic centers the SDSS believes to be two separate galaxies

METHODOLOGY

The first task to be completed in executing this project was the collection of data to use. I began with using Data Release 7[3] from the SDSS Legacy Survey to do the initial experimentation. From the SDSS, I needed several parameters that the pipeline could provide to properly compute the galaxy pair alignment and separate them into the proper categorization. For a galaxy to be considered in this data set, it must have a redshift of $z > 0.001$, a Petrosian r-band apparent magnitude in the range of $14.0 \leq r \leq 17.77$, and a radius that is greater than twice the size of the point spread function radius ($R > 2R_{PSF}$). The value of the radius needing to be over twice the size of the point spread function is due to the need of the galaxies being sufficiently resolved enough to classify when there is a case of the pipeline shredding the object. Also, in the case $R \sim R_{PSF}$ the image of the galaxy is smeared to become more circular, thus making the axis ratio (q) an incorrect value when making calculations. After applying all these needs for parameters, the SDSS gave a data size of 424,276 galaxies.

After the pipeline determined the number of galaxies that fit in the description provided, it then was able to return requested parameters about each galaxy. The first needed was the right ascension and declination. These are comparable to the longitude and latitude coordinates for a galaxy on the night sky. These coordinates, along with the redshift of each galaxy, are then used

[3] Abazajian et al. (2009).

[4] Introduction to Cosmology (2017).

to calculate both the angular separation (θ) and the projected physical separation (r_p) of the two objects with the equations given for this calculation given by:

$$\theta = \sqrt{(\alpha_1 - \alpha_2)^2 \cos(\delta)^2 + (\delta_1 - \delta_2)^2} \quad \text{and} \quad r_p = 1.28\, Mpc * \left(\frac{\theta}{1\, arcmin}\right) * (z_1 + z_2)/2 \quad (1)^4$$

where α is the right ascension on the sky, and δ is the declination (using $H_0 = 68$ $kms^{-1}Mpc^{-1}$ in conversion to projected physical separation within the limit $z << 1$)[5].

For two galaxies to be compared in this experiment, they must have an angular separation $\theta < 3°$ or a projected physical distance $r_p < 30$ Mpc, as galaxies beyond these points have lost any sort of shape alignment correlation with one another[6]. Galaxies within these separations are then checked to ensure that they are not a victim of shredding, as this shredding tends to favor the parallel alignment correlation at low separations due to bright misidentified objects within the galaxy lying within the plane of that galaxy. Just from removing the shredded pairs at the lowest separation resulted in around 200 unusable galaxy samples, with around 600 entries in all that were determined to be sufficiently indistinguishable and thus "bad galaxies" for the analysis.

With distances calculated and errors in the data set removed, the computations for the alignment correlation can begin. To calculate this, the parameters e_x, e_+, q_m, and φ_m are gathered from the SDSS data pipeline. The e_x and e_+ parameters are used to calculate q_{am} and φ_{am} for the adaptive moment model of the galaxy, which is based on a Gaussian weight function to find where the shape best aligns with the galaxy image. These two parameters then become q_{am} and φ_{am} through the equations:

$$q_{am}^2 = (1 - e)/(1 + e) \text{ , where } e^2 = e_+^2 + e_x^2 \text{ , and } tan2\varphi_{am} = e_x/e_+ \qquad (2)^{7\&8}$$

giving values that are comparable to q_m. and φ_m. The values of q_m and φ_m are retrieved from the model calculation of the galaxy alignment, where these parameters are assumed to be uniform throughout the galaxy. This comes from the SDSS comparing characteristics of the galaxy and determining whether it fits better to a de Vaucouleurs[9] or exponential surface brightness profile. These profiles are based on comparison of light intensity as a function of galaxy radius, I(R), where a de Vaucouleurs galaxy light profile has $\log I(R) \propto -R^{1/4}$, and an exponentially spread galaxy light profile has $\log (R) \propto -R$. The data pipeline decides based on likelihood which profile has the better fit and thus chooses that q and φ value to represent the model galaxy alignment calculation.

Now having the axis ratio and position angle found in both the model and adaptive moment fits, we can form a complex shape parameter for each galaxy. This results in a vector with the form:

$$\overrightarrow{\chi} = \frac{1 - q}{1 + q} e^{i2\varphi} \qquad (3)$$

where each q and φ can be substituted in for either model or adaptive moment calculation. From this, the alignment function between two galaxies is calculated by taking the dot product of the two complex shape parameters to give the equation[10]:

[5] Planck Collaboration XIII (2016)
[6] Okuruma et al. (2009)
[7&8] Bernstein & Jarvis (2002, Hirata) & Seljak (2003)
[9] de Vaucouleurs (1948)
[10] Brainard et al. (2009).

$$C_{xx} = \vec{\chi}_1 \cdot \vec{\chi}_2^* = \frac{(1-q_1)(1-q_2)}{(1+q_1)(1+q_2)}[cos2\varphi_1 cos2\varphi_2 + sin2\varphi_1 sin2\varphi_2] \quad (4)$$

When the calculation is made, this function will return a value between +1 and -1, where +1 is a pair of highly elongated galaxies that lie exactly parallel with each other ($\varphi_1 = \varphi_2$) and -1 equates to highly elongated galaxies that lie perpendicular to each other ($\varphi_1 = \varphi_2 \pm 90°$). An example of this calculation is show in Figure 3 where a C_{xx} value of about 0.25 is returned even though visually they appear to lie almost completely parallel to each other.

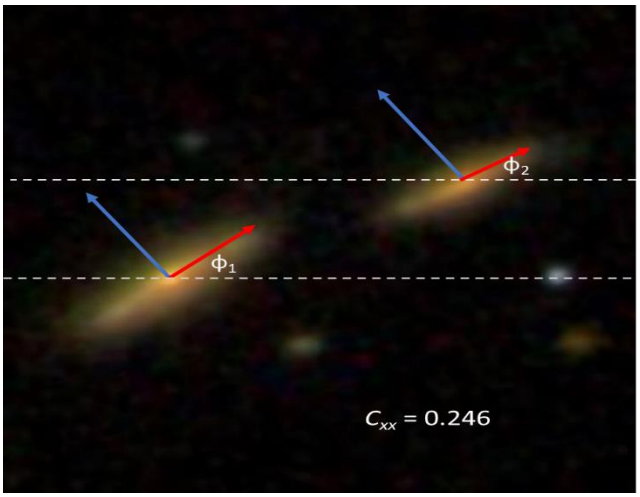


Figure 3 – An example of how the alignment correlation is made where the blue and red lines are the two axis-ratios used to find q and the φ angle is given from the galactic longitude

The alignment correlation function is then calculated for all galaxy pairs within 30 Mpc of projected physical separation and 3° of angular separation. Once the function is applied, the data is put into one of 20 logarithmically sized bins based upon separation. In the initial section of this data analysis, there is no specifying characteristic used to sort the data and a comparison is done with all galaxies within this range. In order to have coherent data every single point is not plotted but rather a weighted average of each logarithmic bin is plotted to show the effects of alignment correlation as a function of separation. The smallest separation bin had 126 galaxy pairs and the largest separation had 30.4 million galaxy pairs.

After running several different variations of the calculation, it was discovered that the adaptive moment data was slightly inconsistent in comparison to the model data. This is due to the point spread function equation used to find the axis ratios and position angles being not sufficiently well resolved in certain cases and calculations not agreeing with the model calculations. To check the accuracy of both types, several hand calculations and visual comparisons were made in order to determine which method holds the correct results. This analysis showed the model data retrieved from the SDSS had more accuracy, and thus for the rest of this paper will be the data that is being referenced unless otherwise stated. With an accurate set of data now confirmed, it is then plotted for both angular and physical separation to begin an analysis of galaxy shape alignment. These data are shown in Figure 4 where we see the

initial trends of parallel at the lowest separations, perpendicular at slightly larger separations, and losing correlation past a certain point.

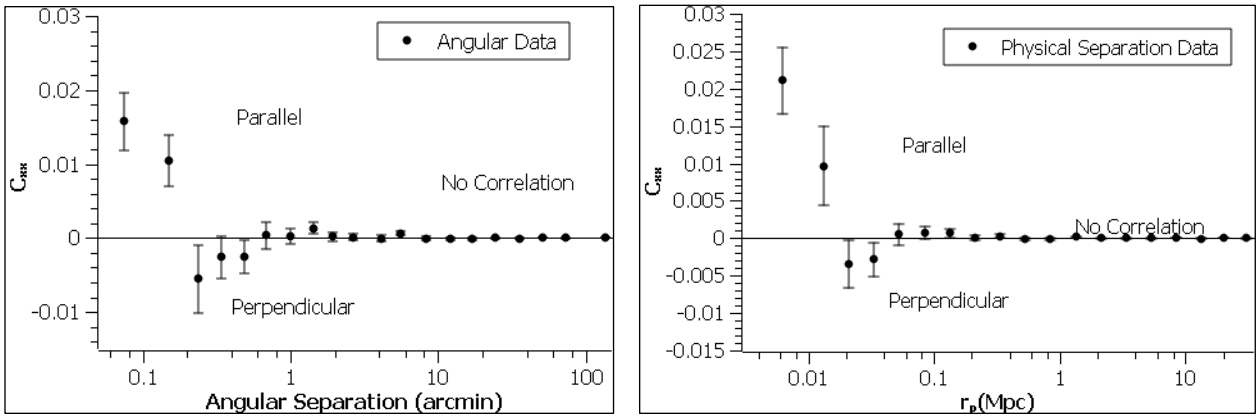


Figure 4 – Data of all galaxies in the data set compared with each other based on their angular separation in arcminutes (left) and projected physical separation in Mpc (right) in which both data sets show the same overall shape with the error bars representing the error in the mean of each bin

From looking at Figure 4, an initial trendline in the data starts to become apparent and it follows the general shape of a gravitational potential. To test this, I ran several different types of provided fits and then an effective gravitational potential to ensure that it is the best fitting provided on the data. It is important to note that this fit is purely coincidental and does not presume the alignments seen are the result of a gravitational potential. The equation being used for the fitting is

$$C_{xx}(x) = \frac{a}{x} + \frac{b}{x^2} \quad (5)$$

where x is the type of separation being measured. Applying this to both graphs, this returned values for the angular separation of a = 4.599e-05 (arcmin), b = 8.712e-05 $(\text{arcmin})^2$, and reduced $\chi^2 = 1.64$. This fit is only used for points up to about 3 arcminutes and 200 kpc as after this the data loses correlation and follows a linear fit across the zero line. The physical separation fit returned values of a = -1.815e-05 (Mpc), b = 9.629e-07 $(\text{Mpc})^2$, and reduced $\chi^2 = 1.58$. These new fits are shown in Figure 5 and will be used in the rest of analysis of graphs for the different types of data splitting. Noting that when the leading term is a negative value, we find that $C_{xx} = 0$ at x = -b/a giving a turning point prediction of when the alignment goes from parallel to perpendicular (in this case x = 53 kpc).

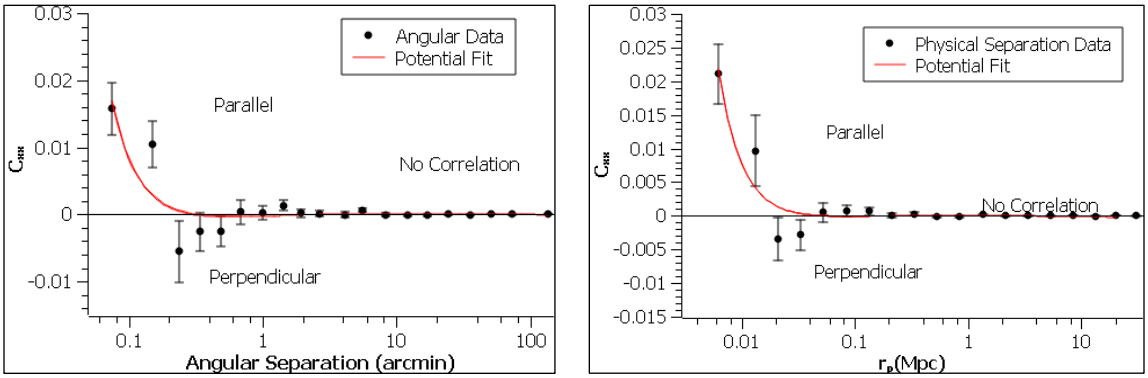


Figure 5 – The $C_{xx}(x)$ fits are now shown on the graphs which shows that the perpendicular points do not have as much of an effect on the fit in comparison to the parallel points

COLOR SEPARATION

To see the effect different galaxy classifications have on the alignment correlation, I split them between early and late type. With there being several ways in which early and late type galaxies can be defined, I began with separating them by color. In general, early type galaxies are associated with being red in color while their late type counterparts are associated with being blue. To make this separation I used the equation[11]

$$(u-r) = 2.294 - 0.146(M_r + 21) - 0.0178(M_r + 21)^2 \qquad (6)$$

where the parameters u and r are the model magnitudes and M_r is the Petrosian magnitude given by the SDSS. This equation provides a separating line based on K- corrected absolute magnitudes where if $(u-r) > 2.294 - 0.146(M_r + 21) - 0.0178(M_r + 21)^2$ then the galaxy is red in color and an early type galaxy. Thus if (u – r) is less than this it is a blue galaxy that is considered late type. This is an updated equation based off the previous Baldry et al (2004)[12] equation used to separate galaxy color by luminosity. With this updated equation, this fit provides an accurate division line shown in Figure 6.

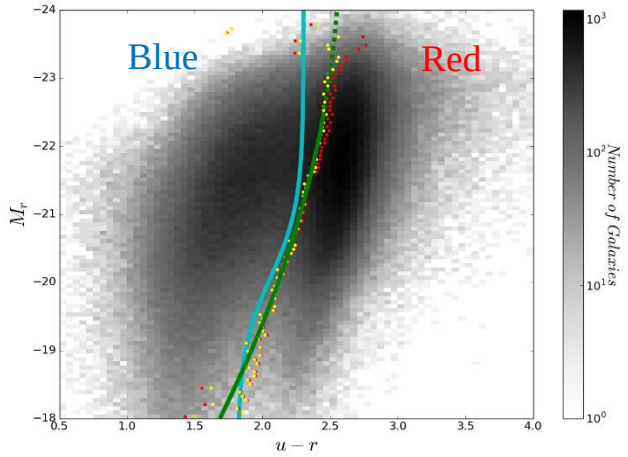


Figure 6 – This shows the Baldry et al. separation in blue and the updated James and Ryden separator in green, which is a quadratic that works to separate the red and blue populations accurately (except at very high luminosities where the red and blue regions begin to converge).

[11] James & Ryden et al. 2017
[12] Baldry et al. 2004

With a distinct separating line made between early and late type galaxies, this can be adapted into the code for testing galaxy alignment. To do this I compared the galaxies in three different requirements. The first was testing only early type galaxy alignment meaning for a pair's alignment to be tested, they both had to be red. This process was then repeated for late type in which both galaxies needed to be blue. The last was comparing early and late pairs together, which tested neighboring galaxies that one sorted into red and the other had to sort into blue. This gave a complete coverage of all possible galaxy pairs based on color separation. The data is then passed through these parameters and separated into logarithmically spaced bins that have been averaged to give the graphs presented in Figure 7. I have chosen to reference the physical separation data within the paper; the angular separation data can be found in the Appendix.

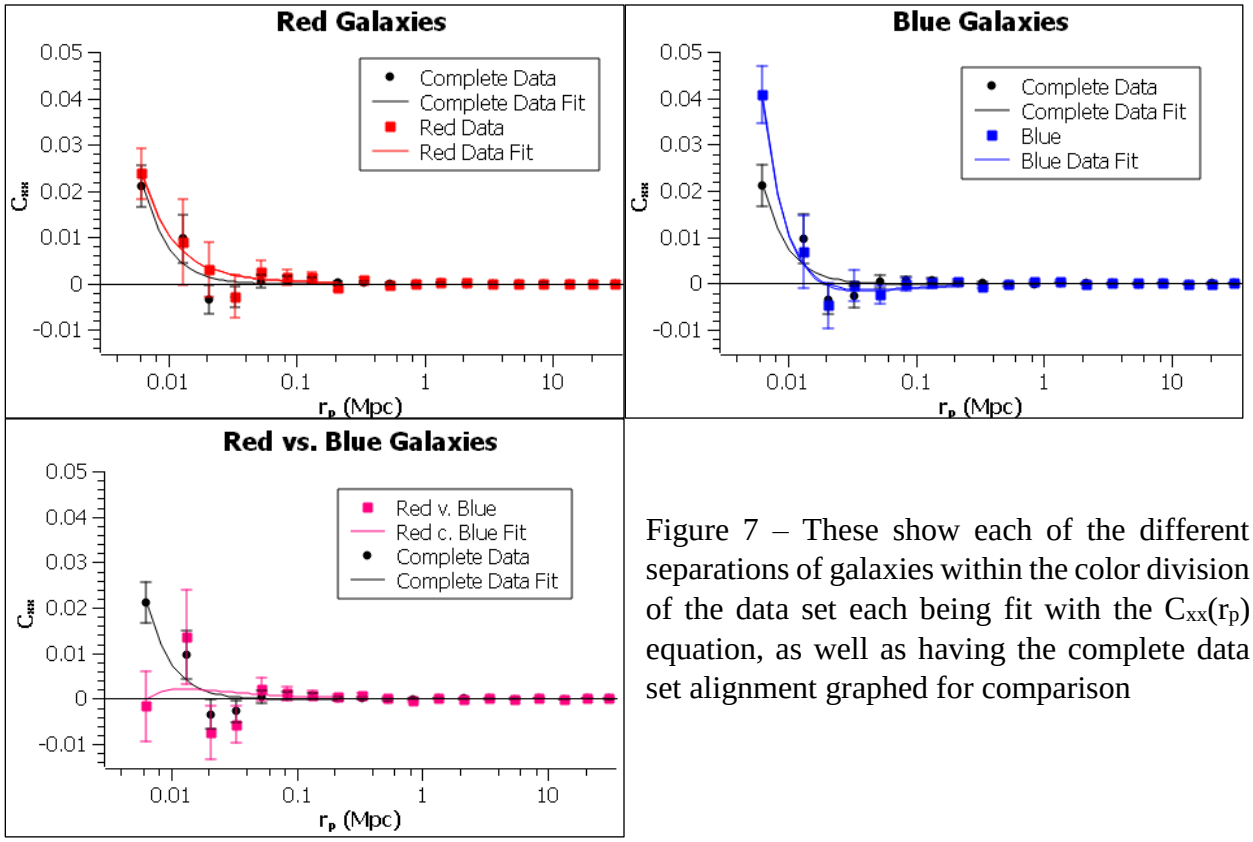


Figure 7 – These show each of the different separations of galaxies within the color division of the data set each being fit with the $C_{xx}(r_p)$ equation, as well as having the complete data set alignment graphed for comparison

When looking at this data we notice that both the red and blue galaxies follow the same trends that the complete data set followed. This trend is that at low separations up to about 20 kpc the pairs tend to align parallel to each other. The blue galaxies have a parallel alignment in the first bin that is about twice that of the complete data, while the red galaxy pairs have an alignment that is almost exact with the complete set. Continuing into higher separations leads to a decrease in parallel alignment until the 20 kpc mark where the galaxies begin to take on a slightly perpendicular alignment. It is the overall trend that the blue galaxy pairs have overall larger amplitudes of alignment in both the parallel and perpendicular regimes, while the red

galaxy pairs follow the same trends with a weaker alignment. This continues until about 100 kpc where red and blue galaxies, along with the complete data set, lose a correlation and center around the zero line. These graphs lead to evidence that overall late type galaxies have higher parallel and perpendicular alignments when compared to their early type counterparts (and will be shown further in the coming sections). These data also show that when comparing a red galaxy against a blue galaxy, they tend to lose any correlation. Two galaxies from different classifications do not have an intrinsic orientation of alignment with their neighbor, at least when looking at this division through color separation.

These data are also fitted by the same function as earlier where the alignment correlation can be calculated more precisely as a function of physical separation. The red galaxy pairs follow the same path of the complete data set in that the parallel aligned points dominate the fit and does not show an area where they would have a perpendicular alignment. The equation for the red galaxy pairs is:

$$C_{xx,Red}(r_p) = \frac{3.751 \times 10^{-5}(Mpc)}{r_p} + \frac{6.854 \times 10^{-7}(Mpc)^2}{r_p^2} \qquad (7)$$

with a reduced $\chi^2 = 1.04$ when fit to the data before the point of no correlation. The same is done to the blue galaxy pairs which follow the same overall trend as the latter but with higher amplitudes. In this fitting there is a strong enough amplitude in the perpendicular alignment for the fit to go below the zero-point line and make predictions of how much perpendicular the alignments will be in this range. The equation for the blue galaxy pairs is:

$$C_{xx,Blue}(r_p) = \frac{-1.116 \times 10^{-4}(Mpc)}{r_p} + \frac{2.265 \times 10^{-6}(Mpc)^2}{r_p^2} \qquad (8)$$

giving a reduced $\chi^2 = 1.002$ for the equation with data containing alignment correlation and a turning point at x = 20.3 kpc. The red vs. blue fit shows the irregularity of the data as its fitting line follows an opposite trend than the previously fit data and is better fit to $C_{xx} = 0$. With this data being lacking any correlation trends I have chosen to omit its data fit for calculations in the main paper and can be found later in the Appendix. Overall this data has shown us the beginnings of a trend in late type galaxy pairs having higher amplitudes of alignment when compared against the weaker aligned early type.

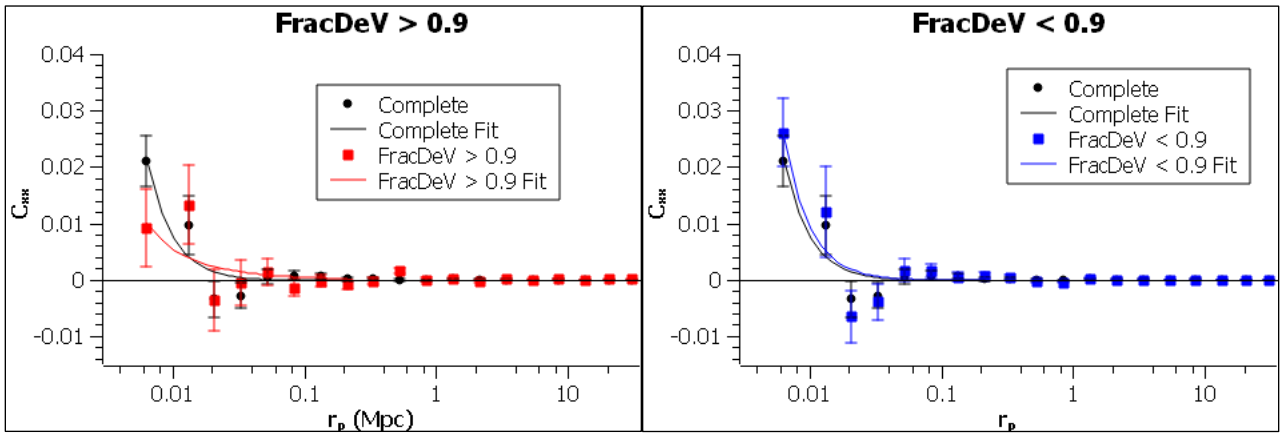
LIGHT PROFILE SEPARATION

The next separation done on the data set is based upon the light profile of the galaxy. This light profile is characterized by how light is distributed in the galaxy and whether this light is centrally concentrated or is exponentially spread throughout. To classify a galaxy's light profile, I used a de Vaucouleurs model. As mentioned previously, a de Vaucouleurs model uses light intensity as a function of radius which is given by the full equation of:

$$I(R) = I_e \exp[-7.67(R/R_e - 1)^{-1/4}] \quad (9)$$

which gives the ratio mentioned before of log I(R) $\propto$ -$R^{1/4}$ for light concentration in this model (where R_e is the effective radius containing half the galaxy's light). The SDSS uses this equation for light profiles and adapts it into their own parameter called fracDeV. This fracDeV parameter that is returned gives a number between 0 and 1 for each galaxy, where 1 equates to a heavily central concentrated light profile galaxy and 0 is a completely exponential spread light profile galaxy[13]. These two classifications can then be used to split the galaxy by early and late type. An early type galaxy tends to have a more centrally concentrated light profile and therefore any galaxy with fracDeV > 0.9 is put into an early classification. A late type galaxy is associated with light being spread throughout the galactic disk and thus a galaxy with fracDeV < 0.9 is put into the late type category.

With this distinction in light profiles made between early and late type, the parameter can be used in the same way previously used in the code for color separation to properly group the galaxy pairs. In this I grouped pairs of galaxies with fracDeV < 0.9 with themselves along with pairs of galaxies both with fracDeV > 0.9. Then once again for completeness I compared the case where one galaxy had an exponential light profile while the other was centrally concentrated to see if there are any statistical trends in this separation of data. The physical separation data is shown in Figure 8 with the angular separation being for comparison in the Appendix.



[13] Unterborn & Ryden 2008

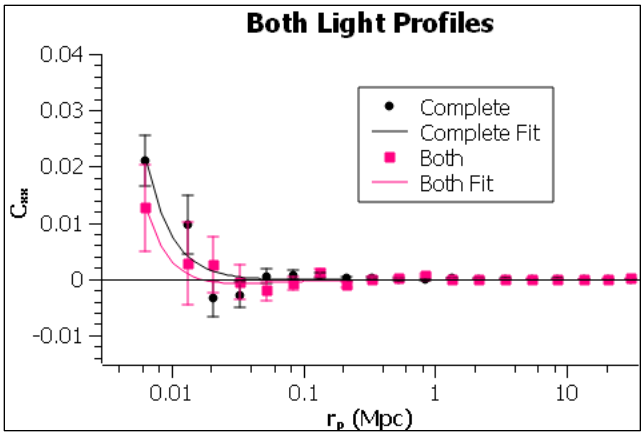


Figure 8 (Including above graphs) – Show the three different separations used by light profile with the complete data set included to use for comparisons

Looking at this new data set, the first noticeable aspect is that all the data follow the same overall trends followed in the complete data set. Once again, we see the continuum of parallel alignment to 20 kpc, perpendicular alignment to 100 kpc, then loss of alignment correlation after this separation. In these data we see that the late type galaxy pairs, fracDeV < 0.9 have almost triple the parallel alignment correlation (0.026) compared to the early type galaxy pairs (0.009) in the smallest separation (~8kpc). Thereafter, we see that both data sets follow similar amplitude alignments as separation increases, both of which following closely the alignment correlation of the first data set. There is still a slightly more perpendicular alignment seen in the fracDeV < 0.9 compared to its counterpart that follows the trend between early and late type, but this difference is not as prevalent in comparison to the color separation previously used. Another big difference seen is in the comparison of pairs with one early and one late type galaxy. Previously it was shown that there is no correlation when the two classifications are directly compared with each other. Now the data follows a trend similar to that of the complete data set with amplitudes that slightly vary from the data. This can imply that there is a correlation present and that the red vs blue alignment has an error or that this alignment correlation only comes when comparing late and early types based on light profiles and comparing the galaxy color adds a characteristic to the galaxies that result in a loss of correlation. This will be further discussed after the comparison in spectroscopy to obtain a better understanding of how one early type and one late type tend to align.

With the data now properly sorted and graphed, it can also be fitted by the $C_{xx}(r_p)$ equation used on the other data sets. Beginning with the fracDeV > 0.9 data, the graph shows this fit is also dominated by the parallel alignment points and does not go below the zero-point line to calculate where perpendicular aligning galaxy pairs will lie as a function of distance. This line is also comparable to the complete data set fit, but the fitting for this line is also thrown off by the second point in the data not being lower than the first thus making it more complicated to fit the decay expected in the first part of the fitting. Because of this we expect a less steep decay in the function similar to the plot of red galaxies and that is seen in:

$$C_{xx,>0.9}(r_p) = \frac{4.517 \times 10^{-5}(Mpc)}{r_p} + \frac{1.349 \times 10^{-7}(Mpc)^2}{r_p^2} \qquad (10)$$

and a reduced χ^2 = 1.29. The fracDeV < 0.9 is then fitted in the same way in which a fit similar to that of blue galaxies is expected. This graph has lower overall amplitude in comparison to the blue pairs of galaxies, but higher amplitude compared to red or the complete data set. Similar to the blue fit, this late type equation goes below the zero-point line, this time only slightly which shows that the parallel galaxy alignments have a stronger influence on the fit. This gives the equation:

$$C_{xx,<0.9}(r_p) = \frac{-1.486 \times 10^{-5}(Mpc)}{r_p} + \frac{1.110 \times 10^{-6}(Mpc)^2}{r_p^2} \qquad (11)$$

Giving a turning point at x = 74.7 kpc and a reduced χ^2 = 1.77 which is higher than the previous late type galaxy equation. In this case where comparing both the types of data does not result in irregularities, I have decided to include an analysis in the results. Instead of having a trendline that follows an opposite pattern than the other data sets, this trendline follows the general shape of the complete data and in turn the other previously graphed data. Even though it does not have as high of alignment amplitudes even in comparison to early type data, most of the points still align with the overall shape of galaxy alignment correlation. This results in the equation:

$$C_{xx,Both}(r_p) = \frac{-5.019 \times 10^{-5}(Mpc)}{r_p} + \frac{8.195 \times 10^{-7}(Mpc)^2}{r_p^2} \qquad (12)$$

resulting in a turning point at x = 16.3 kpc and a reduced χ^2 = 1.28 which is similar to the value for the fracDeV > 0.9 fit. With the complete analysis done on the light profiles, it is seen that overall the data shows the same trends found in the previous color separation in that there is an intrinsic alignment of galaxies that is a function of separation for both early and late type galaxy pairs. It showed the continuation of the trend that late type galaxy pairs have an overall greater alignment in both the parallel and perpendicular regimes. The difference was that in this set there was the first instance of there being a possible alignment correlation between pairs consisting of an early and late type galaxy that was not seen previously, showing this will be an aspect to look more into with the spectroscopic data.

SPECTROSCOPIC SEPARATION

The last separation done on the data was done by grouping the galaxy pairs with respect to their spectroscopic data. This is found through the ratios of hydrogen, ionized Nitrogen, and doubly ionized Oxygen. These ratios and strengths of different emission lines can tell us whether the excitation of that galaxy is dominated by star formation or from an active galactic nuclei (AGN). A galaxy with ongoing star formation is classified as a late type galaxy, while one that is AGN dominated is put into the early type category. This is due to star formation showing that a galaxy is still young enough to be continuously producing stars that give HII emissions opposed to a galaxy that is dominated by its AGN, a supermassive black hole located in the center of the galaxy, which is a quality of galaxies that no longer have an abundant amount of star formation. A composite galaxy is one that is a mix of both types[14&15], where it is seen there is still some ongoing star formation as well as having characteristics of an AGN. The composite galaxy is not considered early or late type, as it can belong to either group depending on the other characteristics of the galaxy. Figure 9, known as a BPT (Baldwin, Phillips, and Terlevich)[16] diagram, shows how these galaxies are separated based upon HII emission and the dividing lines between these different regions. The NII and OIII are emission lines that come from the produced photons during electron transitions when nitrogen and oxygen become ionized. The Hα and Hβ are the lowest energy transitions within the Hydrogen Balmer series, thus the BPT diagram was made to show the logarithmic ratios of these emissions and make comparisons on the areas in which the ionization of these elements is a result of ongoing star formation or is a product of the light concentration around the observed galaxy's AGN.

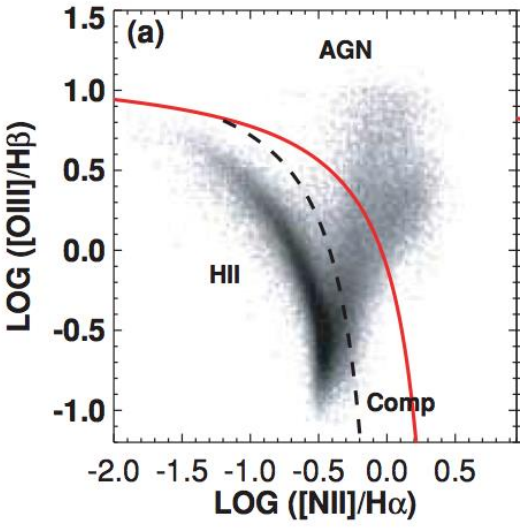


Figure 9 – This graph displays the separations of HII, composite, and AGN regions based upon their ratio of emission line strength.

With the separate regions defined for early (AGN) and late (HII) type these conditions are put into the algorithm to compute alignment correlation. The initial comparison done is of the early and late regions, where pairs of galaxies that are star forming dominated and AGN dominated are grouped together and compared. The pairs of galaxies consisting of one being star forming and one AGN dominated are also compared for completeness and to analyze any correlation between an early and late type pair as done previously. This data is then graphed in Figure 10 with the angular separation measure for distance in the Appendix for comparison.

[14&15] Kewley et al. 2001, Kauffmann et al. 2003
[16] Baldwin, Philips, and Terlevich 1981

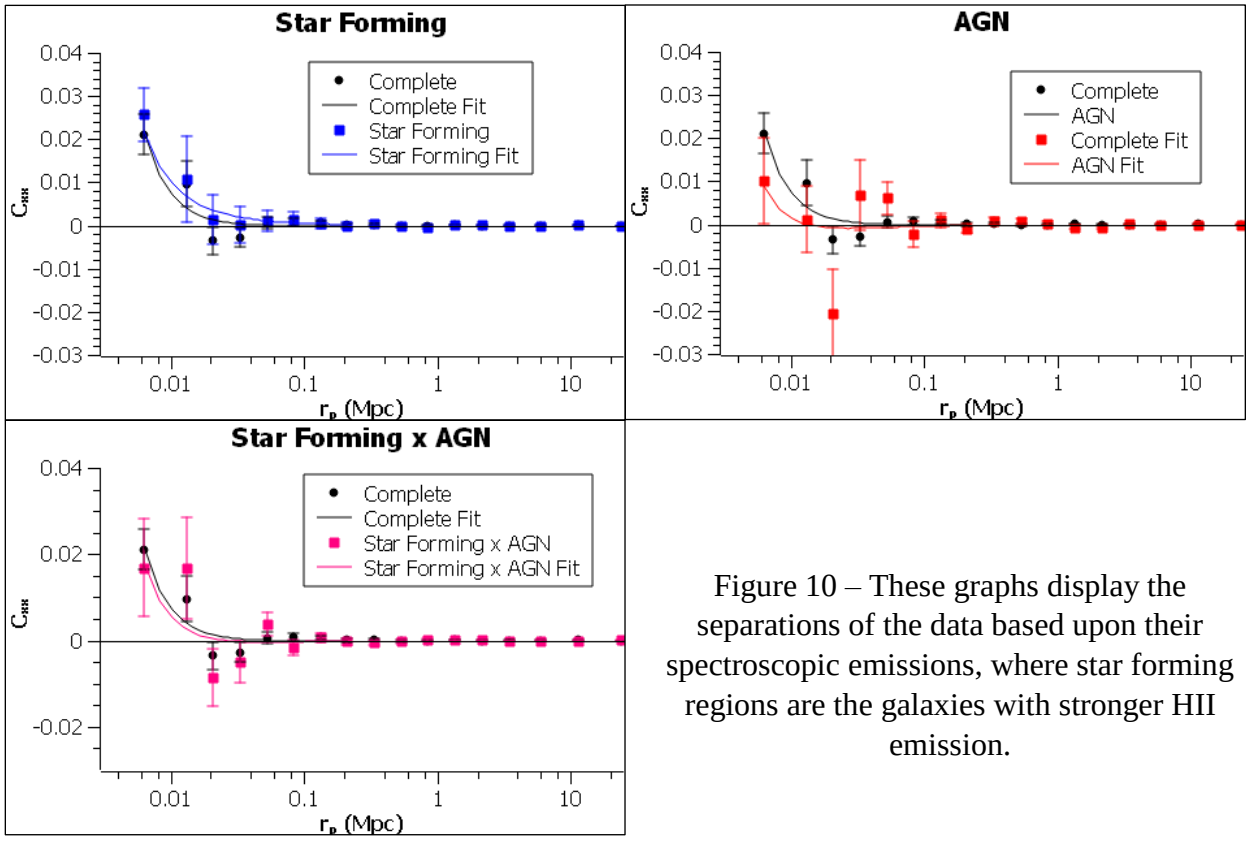


Figure 10 – These graphs display the separations of the data based upon their spectroscopic emissions, where star forming regions are the galaxies with stronger HII emission.

When looking at this data set, there are several trends that are comparable to the graphs in Figure 8. As the previous separations have shown, the late type galaxy pairs continue to have a greater amplitude of parallel alignment in projected physical separations under 20 kpc. In this spectroscopic separation, a change happens where the early type galaxies have one of the highest amplitudes for perpendicular alignment. At the transition from parallel to perpendicular regimes, the early type galaxy pairs have a high perpendicular alignment as opposed to the smooth transitions previously seen in separations. Along with this, these AGN pairs show a high deviation from the complete data set showing that these AGN dominated galaxies have anomalies present due to their central supermassive black holes. The comparison between a star forming and AGN galaxy shows the same trend as the previous early and late type separation used in the de Vaucouleurs model. This follows the same trends as the complete data with high parallel alignments and then small perpendicular alignments up to 100 kpc before losing any alignment correlation. With the spectroscopic data following this trend as well, this shows that the initial comparison between a red and blue galaxy has an irregularity as there was no correlation between the two, while the next comparisons of an early and late type showed the overall statistical trends expected of galaxy alignment.

Now that the initial separation of early and late type has been graphed, these data sets can be fit with the same gravitational potential equation used to quantify alignment as a function of

projected physical separation. Beginning with the early type fit for AGN galaxies, the data returns the equation:

$$C_{xx,AGN}(r_p) = \frac{-4.789 \times 10^{-5}(Mpc)}{r_p} + \frac{6.463 \times 10^{-7}(Mpc)^2}{r_p^2} \qquad (13)$$

where once again noting the negative term in the first part of the equation shows that this fit can account for the calculation of perpendicular galaxies (turning point at x = 13.5 kpc). Even with this negative term, as seen with the other equations, there is still a dominance by the parallel galaxy pairs that make these equations favor them in calculation. This equation gives a reduced $\chi^2 = 1.54$. The star forming galaxy pairs in this set follow the complete data set points closely with only slight deviations to distinguish the two sets. Because of these star forming galaxies having overall weaker amplitude signals in comparison to other late type galaxy pairs, the fitting equation resembles the complete data and is given by:

$$C_{xx,HII}(r_p) = \frac{4.627 \times 10^{-5}(Mpc)}{r_p} + \frac{5.649 \times 10^{-7}(Mpc)^2}{r_p^2} \qquad (14)$$

along with a reduced $\chi^2 = 1.2$. This is smaller than the reduced χ^2 found in the other fits and it is seen by looking at the graph for the star forming data that they do not have much of a perpendicular alignment in any regime and thus the equation can easily fit to calculate the degree of parallel alignment for these galaxies. Figure 9 also shows the first instance of a late type classification not having an equation that crosses the zero-point line. There seems to be a result in the spectroscopic data that shows where the parallel and perpendicular alignments come from, in that AGN domination in galaxy pairs leads to perpendicular alignments, while star formation in galaxy pairs leads to parallel alignments. With still more data to be analyzed, this will be explored further into as more correlations are made prevalent. Having this in mind, the early vs. late type galaxy pairs are then fit in order to investigate further into correlations between the two types. This results in the equation:

$$C_{xx,HII/AGN}(r_p) = \frac{-3.771 \times 10^{-5}(Mpc)}{r_p} + \frac{9.529 \times 10^{-7}(Mpc)^2}{r_p^2} \qquad (15)$$

with a turning point at x = 25.3 kpc and a reduced $\chi^2 = 1.32$. Here this equation has a negative leading term and can therefore calculate the perpendicular alignments of pairs, where it can be seen there is a regime below the zero-point line that has about the same amplitudes as the complete data set. This star forming vs. AGN comparison shows a similar equation that comparing both fracDeV sets had which reinforces that there is an intrinsic alignment to be found when comparing an early and late type galaxy that is simply not seen in a red and blue color division.

Figure 9 also shows that there exists a region in between the star forming and AGN dominated regimes that a galaxy can exist in. As mentioned previously, these composite galaxies are a sort of "mix" between these early and late types, in which they show signs of ongoing star formation while still having an AGN. To have a complete analysis of the spectroscopic data the SDSS provides, I then made comparisons of composite galaxies with themselves to understand any correlations these 'undefined' galaxy types may have. Along with this, I also compared pairs of galaxies that consisted of one composite and one star forming, as well as one composite and one AGN to analyze how these galaxies tend to interact with both early and late type. These three comparisons are shown in Figure 11 using alignment as a function of projected physical separation where the data of alignment as a function of angular separation will be in the Appendix to make comparisons.

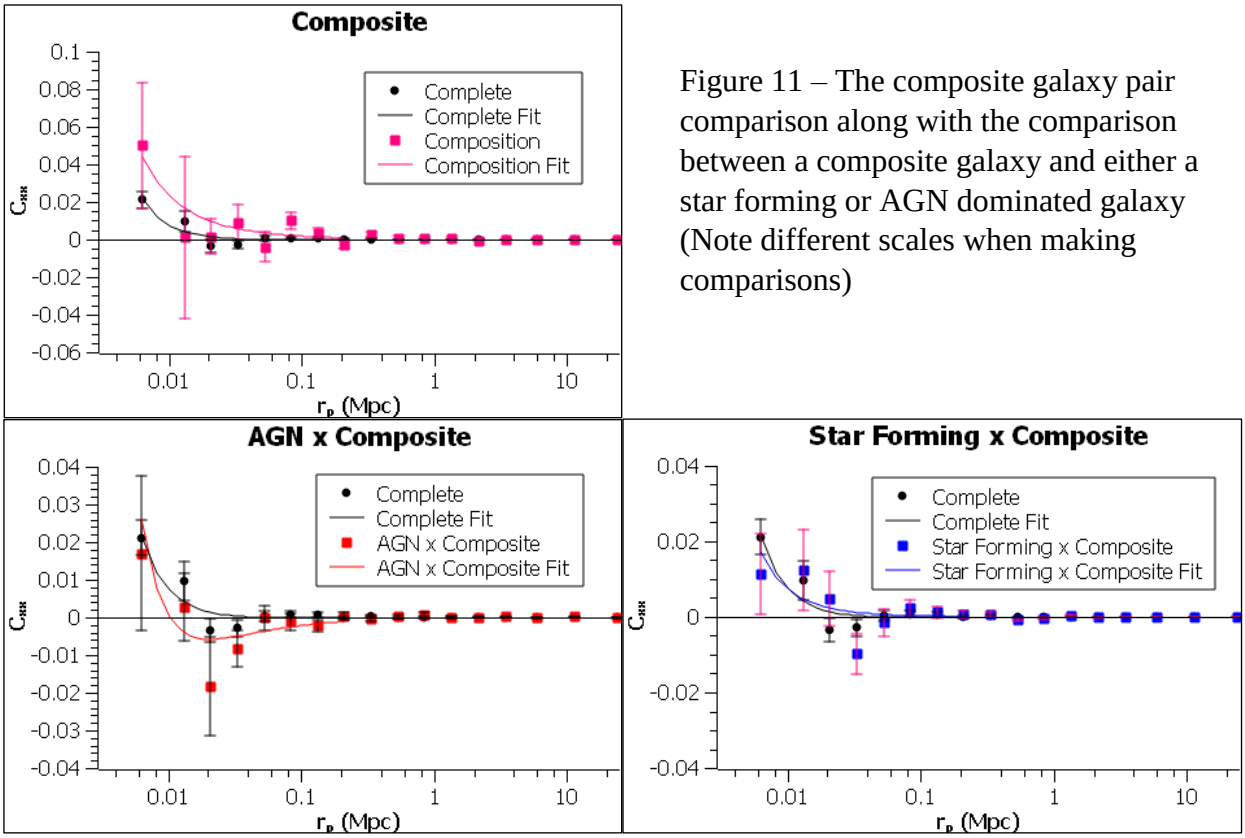


Figure 11 – The composite galaxy pair comparison along with the comparison between a composite galaxy and either a star forming or AGN dominated galaxy (Note different scales when making comparisons)

When looking at the composite comparison, it needs to be noted the difference in the scales between the three graphs. I have left the bottom two with the same scales as the other graphs in order to be able to still see any trends, while the composite galaxy pair graph has scales that are about double the rest. This brings up one of the most interesting findings that these composite galaxy pairs have the highest parallel alignment amplitude out of all the other separations previously used. With this point being the most parallel seen thus far, it cannot be ignored that these points also have the largest associated errors than any other in the data set. This is seen in both two smallest separation bins (<20kpc) which results in the first having a high parallel

alignment and the second having a high perpendicular alignment with the errors considered. Overall, the composite galaxy pairs follow a late type alignment correlation in which it is seen to have the same shape when graphed as the star forming pairs where all the actual data points are found to lie above the zero-point line. In the case of an AGN dominated compared with a composite galaxy, there is now both a strong parallel and perpendicular alignment amplitude in both respective regimes. With the data given thus far this result makes sense as composite galaxies tend to have a parallel influence at the lowest separations (< 20kpc) while AGN dominated galaxies have a perpendicular influence at the slightly higher separations (< 100kpc). The pairs consisting of a star forming galaxy and a composite does not show anything of that much of statistical value. Surprisingly, comparing the alignment correlation between the types that have seemed to show the highest amplitudes in both parallel and perpendicular resulted in data that seems to weakly follow the complete data set and does not provide much insight on how these types influence each other.

Having the initial data on the graph analyzed qualitatively, these are then fit to the gravitation potential equation to begin a quantitative analysis as well. The first fit done is on the pairs consisting of only composite galaxies. Looking at the data presented in Figure 10, this shows this fit has one of the strongest parallel dominations and thus the equation only accounts for parallel aligned galaxies when making calculations as the only imposed perpendicular alignment is in the data point errors. This gives the equation:

$$C_{xx,Comp}(r_p) = \frac{1.706 \times 10^{-4}(Mpc)}{r_p} + \frac{6.608 \times 10^{-7}(Mpc)^2}{r_p^2} \qquad (16)$$

and a reduced $\chi^2 = 1.33$. Since there is a wide range of error attached to these data points the fitting does not have a hard time finding a line that can account for all the data in the set, but the accuracy of this equation must be taken into consideration for this same reason. After completing the composite fitting, the next data set analyzed is the pairs of one composite and one AGN dominated galaxy. Looking at this fit, one of its most notable aspects is the amount of the line below the zero-point transition showing that this equation will be able to better calculate the galaxies in the perpendicular regime. This equation is:

$$C_{xx,CompxAGN}(r_p) = \frac{-2.24 \times 10^{-4}(Mpc)}{r_p} + \frac{2.534 \times 10^{-6}(Mpc)^2}{r_p^2} \qquad (17)$$

resulting in a turning point at x = 11.3 kpc and a reduced $\chi^2 = 1.11$. From this it is noticed that the equation does pass through all the error bars quite well and the negative sign on the leading term shows that the fit is accounting for the perpendicular alignment galaxy pairs. The final fit done was to the pairs consisting of a star forming and composite galaxy. As previously mentioned, this graph does not appear to have a lot of statistical significance with it and does not have a great reliance on perpendicularly aligned galaxy pairs. The fit overall takes the same general shape as the complete data set, only seeming to deviate slightly towards the beginning in the end where the complete data set had a stronger parallel and stronger perpendicular alignment correlation respectively. This fit gives an equation for the line of:

$$C_{xx,CompxStar}(r_p) = \frac{2.318 \times 10^{-5}(Mpc)}{r_p} + \frac{5.313 \times 10^{-7}(Mpc)^2}{r_p^2} \quad (18)$$

with a reduced $\chi^2 = 1.26$. From analyzing the data with the fits attached it is seen that the composite galaxies add a highly varying aspect, not only when compared solely in groups of themselves, but also when comparing with another star forming or AGN dominated galaxy. Because these composite galaxies are a pseudo mix between both an early and late type, it adds a complexity to analyzing the data in deciding what category it best belongs to. In the case of composite galaxies, overall it is important to note that the data seen when comparisons involve them has high errors attached as well as highly varying, even if the graphs show some fits that do go well with the data and trends among the sets.

RESULTS/CONCLUSIONS

Going through this process of comparing the shape alignments of several different galaxy types, we have seen an intrinsic correlation overall on the ways in which early and late type align. As seen through every alignment correlation including the complete data set, galaxies align parallel up to about 20 kpc, then become perpendicular up to around 100 kpc before losing any alignment correlation in the pairs. This data has shown throughout that the late type galaxies, composed of blue in color, having a fracDeV < 0.9, or being star forming dominated resulted overall in higher amplitudes of alignment correlation in both the parallel and perpendicular regimes. The early type galaxies, consisting of red in color, fracDeV > 0.9, or AGN dominated follow the same trends as their late type counterparts but only with lower respective alignment correlation amplitudes. Then in the case of the composite galaxies that lie somewhere between an early and late type, the data was slightly irregular in its trends in that it had the highest individual parallel and perpendicular alignments seen but also the highest errors associated with its data. Along with this, when comparing a composite galaxy with an early or late type resulted in one set that gave favorable results (AGN x Composite) and one that gave results lacking in statistical significance (Star Forming x Composite). Thus, reinforcing the point that comparisons using these composite type galaxies need to take the high errors into account and this is a region that can be investigated into further.

When taking all the data fits and comparing them based on their classification of early and late type, this resulted in two equations:

$$C_{xx,Early}(r_p) = \frac{1.16\times10^{-5}(Mpc)}{r_p} + \frac{4.89\times10^{-7}(Mpc)^2}{r_p^2}$$

$$C_{xx,Late}(r_p) = \frac{-2.69\times10^{-5}(Mpc)}{r_p} + \frac{1.32\times10^{-6}(Mpc)^2}{r_p^2} \quad (19)$$

when each of the coefficients are averaged between the early and late type pairs. The late type equation shows that overall the late type fits have enough influence from the perpendicularly aligned galaxies to have a negative leading term (giving a turning point at x = 49 kpc) in the

calculation while early type does not have the amplitude in the perpendicular alignments to cause this.

Even with the analysis done through this report, there is still many more aspects that can be delved into for a further understanding on galaxy shape alignment. Understanding why galaxy pairs have their intrinsic alignment, whether it's a result of their dark matter halos interacting or maybe early interactions from when the galaxies were first forming, it is important to understand how and why galaxies align with each other to gain insight on the galaxy's formation and future evolutionary track. With the results given in this paper, this serves as a starting point to understanding galaxy shape alignment and a place to begin for observation and analysis in projects to come.

REFERENCES

Abazajian, K., et al. 2009, Astrophysical Journal Supplement, vol. 182, pp. 543

Baldry, I. K. et al. 2004, The Astrophysical Journal, vol. 600, pp. 681-694

Baldwin, J. A., Phillips, M and Terlevich, R. 1981, Publications of the Astronomical Society of the Pacific, volume 93, pp. 5-19

Blanton et al. (2017). *Sloan Digital Sky Survey IV: Mapping the Milky Way, Nearby Galaxies, and the Distant Universe*. http://adsabs.harvard.edu/abs/2017arXiv170300052B

Bernstein, G. M., & Jarvis, M. 2002, The Astrophysical Journal, vol 123, pp. 583

Brainard et al. 2009, *Large-Scale Intrinsic Alignment of Galaxy Images. http://adsabs.harvard.edu/abs/2009arXiv0904.3095B*

de Vaucouleurs, G. 1948, Annales d'Astrophysique, vol. 11, pp. 247

Hirata, C. & Seljak, U. 2003, MNRAS, vol. 343, pp. 459

James & Ryden. (2017). *The Green Valley: Separating Galaxy Population in Color-Magnitude Space.*

Kauffmann, G. et al. 2004, Monthly Notices of the Royal Astronomical Society, volume 346, pp. 1055-1077

Kewley, L. J. et al. 2001, The Astrophysical Journal, volume 556, pp. 121-140

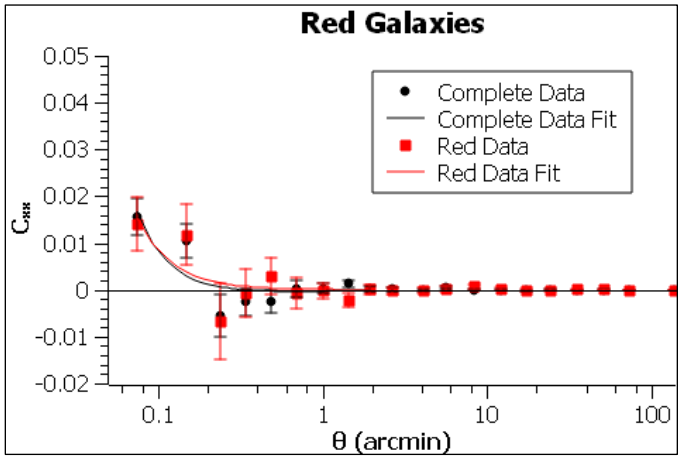
Planck Collaboration, 2016, Astronomy & Astrophysics 594, id. A13, pp. 63

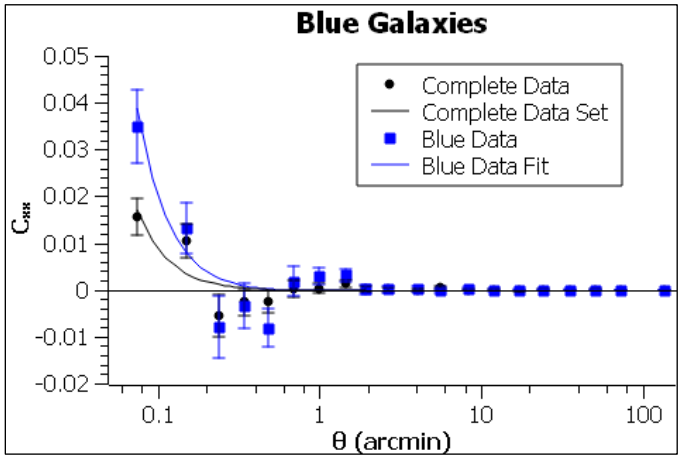
Okumura, T., Jing, Y. P., & Li, C. 2009, Astrophysical Journal Supplement, vol. 694, pp. 214

Ryden, B., *Introduction to Cosmology*, Second Edition (2017), Cambridge University Press, 9781107154834

Unterborn, C. T. and Ryden, B. S. 2008, The Astrophysical Journal, volume 687, pages 976-985, November 2008

APPENDIX





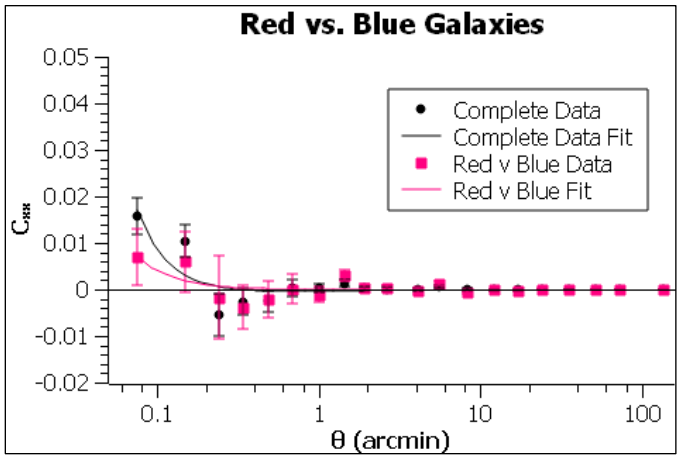


Figure 1A – Graphs of the red (early), blue (late), and red vs. blue galaxy pairs as a function of angular separation each with their own line equations:

$$C_{xx,Red}(\theta) = \frac{-2.1 \times 10^{-4}(arcmin)}{r_p} + \frac{1.005 \times 10^{-4}(arcmin)^2}{r_p^2}$$

with reduced $\chi^2 = 1.23$

$$C_{xx,Blue}(\theta) = \frac{6.892 \times 10^{-5}(arcmin)}{r_p} + \frac{1.813 \times 10^{-4}(arcmin)^2}{r_p^2}$$

with reduced $\chi^2 = 1.45$

$$C_{xx,Both}(\theta) = \frac{-1.85 \times 10^{-5}(arcmin)}{r_p} + \frac{4.197 \times 10^{-5}(arcmin)^2}{r_p^2}$$

with reduced $\chi^2 = 1.15$

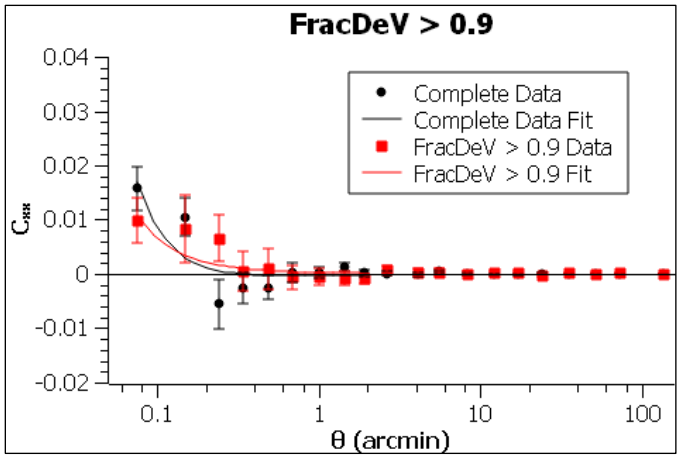
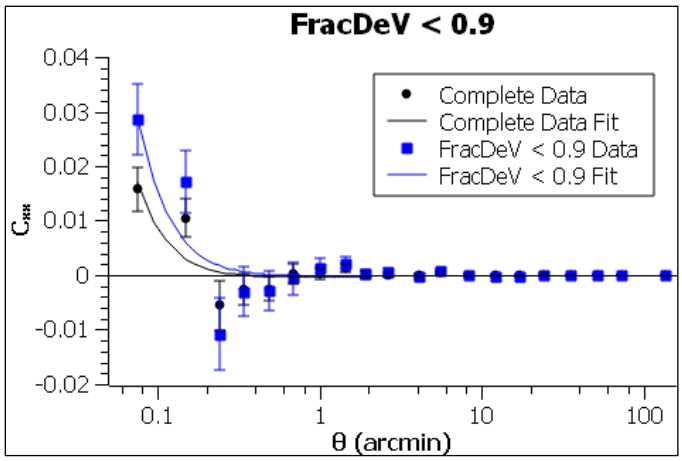


Figure 2A – Graphs of the de Vaucouleurs separations of light profiles for fracDeV > 0.9 (early), fracDeV < 0.9 (late), and the mix of both galaxy pairs as a function of angular separation each with their own line equations:

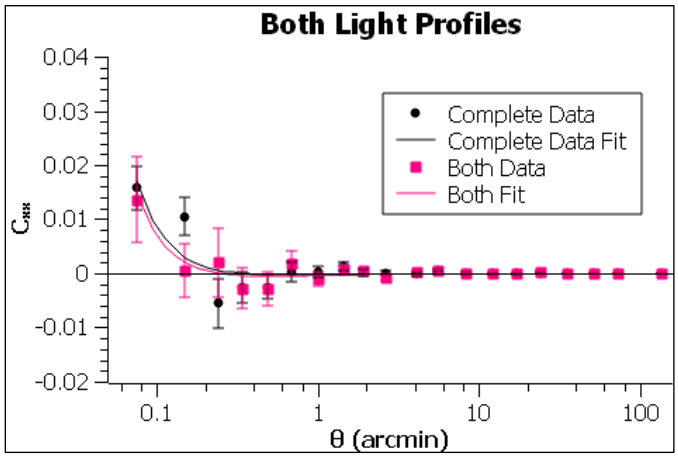
$$C_{xx,>0.9}(\theta) = \frac{2.129 \times 10^{-4}(arcmin)}{r_p} + \frac{4.276 \times 10^{-5}(arcmin)^2}{r_p^2}$$

with reduced χ^2 = 1.1



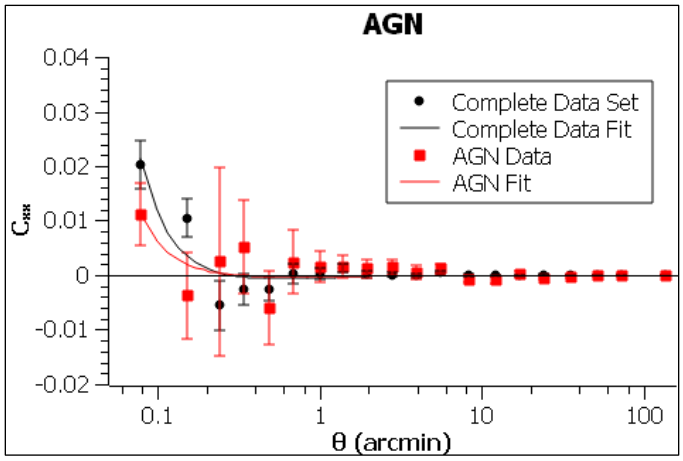
$$C_{xx,<0.9}(\theta) = \frac{-3.45 \times 10^{-4}(arcmin)}{r_p} + \frac{2.014 \times 10^{-4}(arcmin)^2}{r_p^2}$$

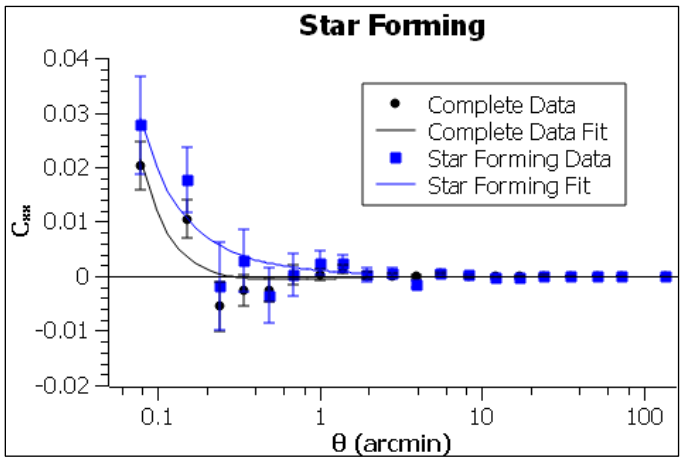
with reduced χ^2 = 1.29



$$C_{xx,Both}(\theta) = \frac{-4.76 \times 10^{-4}(arcmin)}{r_p} + \frac{1.156 \times 10^{-4}(arcmin)^2}{r_p^2}$$

with reduced χ^2 = 1.13





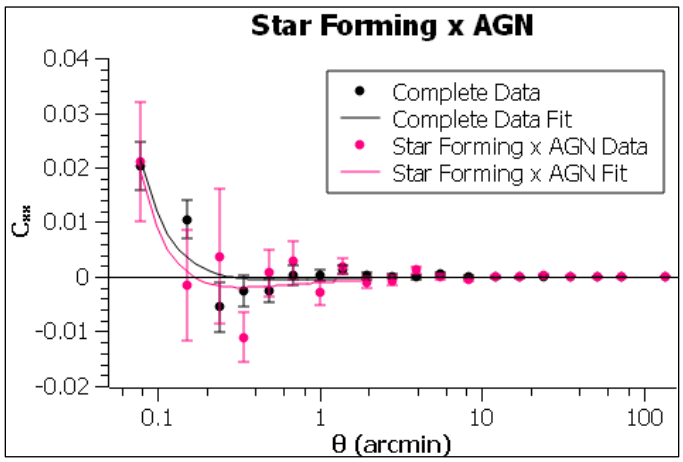


Figure 3A – Graphs of the spectroscopic separations consisting of the AGN, star forming, and AGN vs. star forming galaxy pairs as a function of angular separation each with their own line equations:

$$C_{xx,AGN}(\theta) = \frac{-2.831 \times 10^{-4}(arcmin)}{r_p} + \frac{8.801 \times 10^{-5}(arcmin)^2}{r_p^2}$$

with reduced $\chi^2 = 1.13$

$$C_{xx,Star}(\theta) = \frac{9.455 \times 10^{-4}(arcmin)}{r_p} + \frac{9.924 \times 10^{-5}(arcmin)^2}{r_p^2}$$

with reduced $\chi^2 = 1.33$

$$C_{xx,Both}(\theta) = \frac{-1.297 \times 10^{-4}(arcmin)}{r_p} + \frac{2.14 \times 10^{-4}(arcmin)^2}{r_p^2}$$

with reduced $\chi^2 = 1.1$

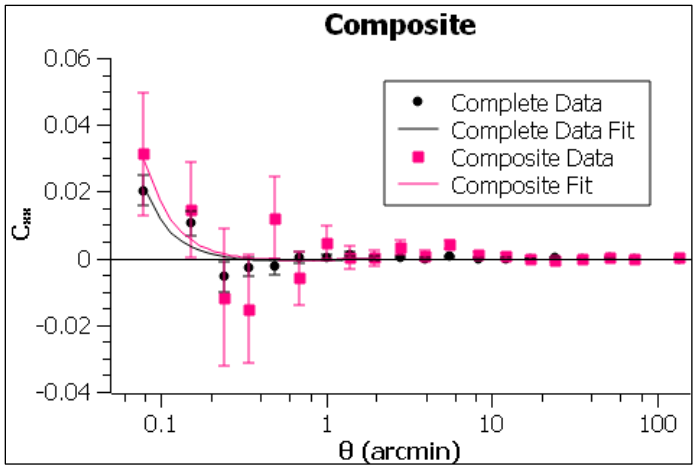
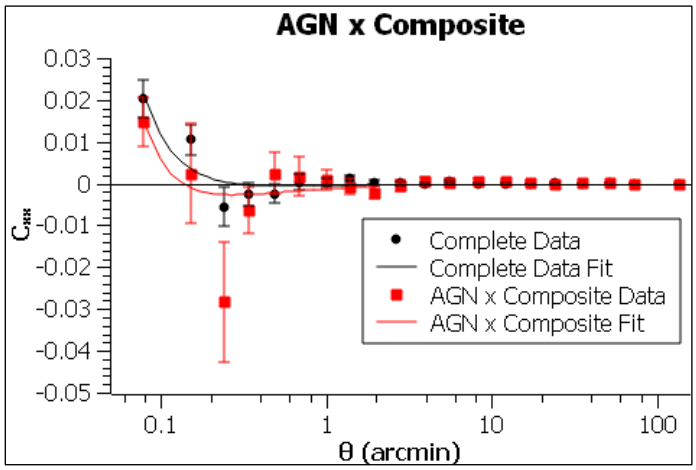


Figure 4A – Graphs of the spectroscopic separations consisting of the composite galaxy pairs, as well as comparing a composite with an AGN galaxy or star forming galaxy. These return equations:

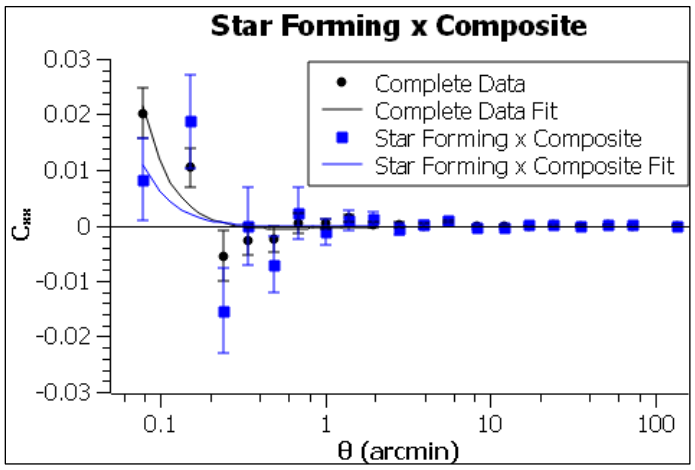
$$C_{xx,Comp}(\theta) = \frac{-6.46 \times 10^{-4}(arcmin)}{r_p} + \frac{2.28 \times 10^{-4}(arcmin)^2}{r_p^2}$$

with reduced $\chi^2 = 1.1$



$$C_{xx,CxA}(\theta) = \frac{-1.48 \times 10^{-4}(arcmin)}{r_p} + \frac{2.02 \times 10^{-4}(arcmin)^2}{r_p^2}$$

with reduced $\chi^2 = 1.3$



$$C_{xx,CxS}(\theta) = \frac{3.382 \times 10^{-4}(arcmin)}{r_p} + \frac{4.423 \times 10^{-5}(arcmin)^2}{r_p^2}$$

with reduced $\chi^2 = 1.4$